

**RESPONSES TO EPA REGION VIII REVIEW COMMENTS ON THE
IDENTIFICATION OF CHEMICALS OF CONCERN
TECHNICAL MEMORANDUM NO. 4 (SEPTEMBER 23, 1994)**

January 16, 1995

These detailed responses are supplied for the purposes of addressing the Environmental Protection Agency's formal comments regarding Identification of Chemicals of Concern Technical Memorandum No. 4 (September 23, 1994) (TM 4). EPA's Comments on TM 4 are in **BOLD** and are preceded by "Comment." DOE's responses to the comments are preceded by "Response."

GENERAL COMMENTS

EPA Noted the following deviations from the standard process for selecting contaminants of concern:

Comment 1, Section 2, Page 27. All data collected under the operable unit 3 (OU 3) field sampling program should be considered when selecting contaminants of concern (COCs). Potential exposure pathways should not be used to limit the data sets under consideration. Subsurface sediments in Standley Lake and Mower Reservoir were excluded incorrectly in DOE's analysis.

Response to Comment 1: All chemical and radiochemical data collected under the OU 3 field sampling program, as well as supplemental radionuclide data (e.g., Jefferson County soils and Great Western Reservoir sediments), were considered initially as the candidate population of data for selection of COCs. No analytical data were disregarded. On February 14, 1994, all parties (including EPA, CDPHE, and DOE) agreed that if sediment core data are not associated with an exposure pathway, the data do not need to be compared to background data for the risk assessment. Therefore, subsurface sediments in Great Western Reservoir were included in the COC identification process because of the possibility (though unlikely) that the reservoir may be drained and could be converted to residential, recreational, or commercial/industrial land uses. The probability of Standley Lake and Mower Reservoir being drained is considered remote. On this basis, subsurface sediments were considered part of the exposure pathway for Great Western Reservoir, but not for Standley Lake and Mower Reservoir. The nature and extent section of the RFI/RI report will present all analytical data regardless of whether the media is of concern in the Human Health Risk Assessment (HHRA).

By including subsurface reservoir sediments in the HHRA process at all, DOE has erred on the side of safety. This is because, in addition to a low likelihood of the reservoir being drained and zoned to allow human contact, it is very unlikely that subsurface contamination would ever be expressed at the concentrations observed in the sediment borings. Any construction activity necessary for development (e.g., clearing, grading, site preparation, excavation, etc.) would remove or extensively homogenize the comparatively small mass of subsurface material so that it would be indistinguishable from background.

All other data sets, with the exception of subsurface soils, were evaluated in TM 4. Surface soil in IHSS 199, surface sediments and surface water in IHSSs 200-202, and groundwater were included in the COC identification process.

Comment 2, Section 3, Page 1. The COC selection process as described on this page and illustrated in Figure 3-1 deviates from the process accepted by EPA, CDPHE, and DOE in three ways:

Response to Comment 2: As a result of technical inconsistency with the assumptions that underlie the approved process (REF strawman; Gilbert, 1993) an alternative approach was used for comparing site data to background data. The alternative approach is referred to as the "weight-of-evidence (WOE) evaluation," since it relies on a series of data evaluation steps and involves the use of professional scientific judgement. The WOE evaluation involves the application of a variety of data analysis techniques in lieu of the formal, quantitative statistical tests recommended by Gilbert. These techniques correspond with the EPA-accepted professional judgement techniques (i.e., spatial analysis, temporal analysis).

DOE discussed the uncertainties and limitations of strictly following the approved flowchart with EPA and CDPHE on March 10, 1994 with the focus on reservoir sediment and surface water data incompatibilities in the approved process. In the ensuing work period, it became apparent that similar issues of technical inconsistency with the approved process emerged with the stream sediments, stream surface waters and groundwater. These issues were discussed in the May 3, 1994 meeting with EPA, CDPHE, and DOE. DOE implemented the WOE evaluation approach as an alternative to the approved process for those media as well. The analyses steps performed in TM 4 are consistent with Region VIII Guidance (October, 1994) as well as national guidance (See Section 5.7, EPA, 1989). Clearly, DOE would prefer a higher level of statistical work in the analysis. However, as was pointed out in the document, in some instances, assumptions and other criteria in the process could not be met; therefore, the alternative within the envelope of guidance was taken (the WOE evaluation).

The WOE alternative was discussed with EPA and CDPHE, and their input sought in meetings on March 10, 1994 and again on May 3, 1994.

Comment 2a. On March 10, 1994, all parties agreed that the Gilbert methodology was not appropriate for reservoir sediments due to the physical differences between on site stream sediments and the off site reservoir sediments. There was no such agreement for the other environmental media within OU 3. We agreed that a weight of evidence approach could be used to address the question of whether or not metals and radionuclides are above background levels in the reservoir sediments. This approach was to be conducted as a first step in the COC selection in accordance with the accepted methodology. Instead, DOE conducted this analysis at the end of the process. The effect of manipulating the process is that chemicals which appear to contribute the largest proportion of the risk within the OU are later explained away as representing background conditions. The true anthropogenic risk drivers may not have been identified.

Response to Comment 2a: On May 3, 1994, DOE presented the COC selection approach used for TM 4 to EPA and CDPHE. It was agreed to by all parties at the meeting that the main OU 3 concern was plutonium and americium and that the level of effort associated with determining if metals are COCs could become disproportionate. Historically, and as early as August 1992, when OU 3 was used as the basis for a "Risk Assessment" seminar with the public, EPA, CDPHE, and DOE recognized that actinide contamination was the focal point of the study. At that time, and throughout the work plan development process, plutonium in soils was a main concern. The assessment of metals was regarded as lesser a concern. On this basis, a statistically based soils sampling program (which ultimately proved compatible with the approved COC process) was specified for the soils (IHSS 199). A sampling strategy to confirm that metals from the Rocky Flats Environmental Technology Site (RFETS) had not impacted the sediments and surface waters using some biased locations in the streams and reliance on existing reservoir sediment plutonium data was developed for metals and radionuclides in sediments, surface waters, and groundwater in the Work Plan (DOE, 1992). Confirmation that these media were not impacted would be inferred if detected concentrations appeared in accord with natural conditions. Thus, from the planning stage to present, it was never envisioned that rigorous statistical methods would be applied on media other than the soils (IHSS 199). The Work Plan was developed in consultation with EPA, CDPHE as well as stakeholders from the Technical Review Group (TRG) and was approved by EPA and CDPHE in 1992.

The following table summarizes the reasons why the compensatory WOE evaluations were necessary for the reservoir sediment, stream sediment, reservoir surface water, stream surface water, and groundwater data.

Table 1 Reasons for the Weight-of-Evidence Evaluation		
Medium	Reason(s)	Discussion
Reservoir sediment (All IHSSs)	No comparable background data set	The Background Geochemical Characterization Report (BGCR) does not contain sediment data from background reservoirs, lakes, or ponds. No other data sets from reservoirs along the front range were found with appreciable sample size. Although other OUs used background seep data from the BGCR, there is no evidence to support that the seep data is comparable to the OU 3 reservoir data.
Stream sediment: IHSS 200: 8 samples IHSS 201: 14 samples IHSS 202: 4 samples Stream surface water: IHSS 200: 4 total/1 dissolved IHSS 201: 4 total/2 dissolved IHSS 202: 0 Groundwater: IHSS 200: 1 well sampled 8 times, repeat samples. IHSS 201: 1 well sampled 8 times, repeat samples.	1. Too few OU 3 samples 2. Disproportionate sample sizes Background Data from the BGCR: Stream Sediments: 20-60 Stream Surface Water: 100 Groundwater: 49 wells (157 samples)	Preliminary statistical evaluations using the approved approach indicated that: 1. Satisfactory confidence and power in the inferential rigorous statistical tests was not possible because of the small sample sizes in the confirmation sampling approach. 2. Rigorous inferential statistical results could not be obtained with confidence owing to disproportionate sample sizes between the OU 3 and background data sets.
Reservoir surface water	No comparable background data set	The Background Geochemical Characterization Report does not contain surface water data from background reservoirs, lakes, or ponds. No other data sets from reservoirs along the front range were found with adequate sample size.

Relevant guidance (EPA, 1994; EPA, 1989; Gilbert, 1993) establish that in RFI/RI assessments chemical concentrations that are indistinguishable from background can be eliminated as COCs from the risk assessment (see Fig. 1 of Region VIII Guidance on COC identification (EPA, 1994). Therefore, focusing the OU 3 assessment on those compounds which can be distinguished from background is consistent to the guidance listed above.

The common sense WOE methodology, in fact, stems in part from Dr. Gilbert's original July 1993 report in which he discussed the use of professional judgment and Geochemical analysis as a significant part of his recommended approach (See Phase V discussion, Gilbert, 1993). In the same paper, Dr. Gilbert emphasized visual data presentations and their interpretations within the site specific setting. The Hi-Lo bar graphs and probability plots

(PROBLOT), which are fundamental tools of the WOE evaluation method, are examples of visual data presentations (Comparing Hi-Lo bar graphs and drawing deductions about means, ranges and variations is analogous to comparing box and whisker plots for the same purpose). Probability plots were cited by Dr. Gilbert and were approved by EPA in the "Straw Man" approach (REF).

Originally, DOE proposed to perform the WOE evaluation as the first step of the identification process--as part of the Gilbert "Tool Box" (as presented in the strawman, EG&G, 1994a). Applying the WOE evaluation early in the process would have screened out many chemicals. Additionally, much of the WOE evaluation is part of the nature and extent of contamination evaluation. The COC selection approach (CPDHE/EPA/DOE, 1994) places a nature and extent of contamination evaluation following the COC selection steps. Therefore, in order to be consistent with this approach (CPDHE/EPA/DOE, 1994) (see Figure 1-1 in TM 4), the WOE evaluation was moved to the last step in the process. This approach adds more conservatism to the process by first applying the toxicity screen and allowing more attention to be focused on the potential risk drivers in the WOE evaluation. If the WOE evaluation is conducted first, the Potential Chemicals of Concern (PCOCs) that continue through the remainder of the COC selection process would be the same PCOCs that went through the CPDHE conservative screen (DOE, 1994b).

Comment 2b. A comparison of maximum chemical concentrations to corresponding preliminary remediation goals (PRGs) is meant to retain those chemicals which are present above the PRG, not to further exclude chemicals which have been identified as contributing the significant portion of the operable unit risk as a result of the concentration toxicity screen. DOE used the PRG comparison incorrectly in OU 3.

Response to Comment 2b: Those chemicals exhibiting maximum concentrations greater than the most conservative PRG (with respect to exposure route (oral or ingestion) and toxicity (carcinogenic or noncarcniogenic) were retained as PCOCs. The PRGs used are included in Attachment 1 to Appendix E. The PRGs are based on residential exposure parameters, with the exception of subsurface sediments which are based on office worker exposure parameters, and use a target risk of 1×10^{-6} or a hazard index equal to 1.0. Comparing the PRG screen results presented in Appendix D (Tables D-1 through D-9) and in Appendix E (Tables E-1 through E-9) with the final COCs in each medium (see Sections 4, 5, 6, and 7) illustrates that no chemicals with concentrations exceeding a PRG were deleted from the COC selection

process at this stage. Only beryllium in IHSS 201 surface sediments failed the concentration-toxicity screen and has a concentration greater than the PRG (1.6 mg/kg vs a PRG of 0.15 mg/kg). DOE's application of the PRG screen clearly achieves the intent of the EPA Region VIII COC Identification guidance (EPA, 1994).

All PCOCs were then subjected to the WOE evaluation for comparison to background levels. It is appropriate to apply the WOE to these PCOCs because some naturally occurring compounds such as arsenic exist in nature at concentrations greater than their respective PRGs.

Regardless of the order of when the PRG screen and background comparisons are performed, those chemicals which can be differentiated from the background per the WOE evaluation and those chemicals contributing a significant portion of the potential risks per the concentration-toxicity and PRG screens will be selected as COCs.

Comment 2c. The accepted statistical methodology for comparing remedial investigation data to background data, the "Gilbert Methodology", was not used for stream surface water, stream sediment, and groundwater. The reasons cited were "insufficient sample size and lack of a comparable data set." We believe the background Geochemical characterization data set is comparable and that it is possible that a statistical comparison can be conducted for these media although the power of the test may not be optimal.

Response to Comment 2c: There are at least four samples for most media by IHSS (see Table 1), and it is possible, mathematically, to perform the Gilbert statistical tests for comparison to background with so few samples and the lack of comparable data sets. However, the uncertainty introduced in the outcome of the statistical tests is likely greater than the approach used in the WOE evaluation. The WOE approach tries to use a variety of information rather than binary hypothesis tests (i.e., OU 3 concentrations greater than background or OU 3 concentrations less than background) than may or may not accurately reflect conditions at OU 3. Statistical analysis on data with so few data points would require additional confirmation. That confirmation was performed using the WOE evaluation.

The issue of whether the background and OU 3 stream surface water, stream sediment, and groundwater data are comparable is not wholly a statistical argument. This issue was discussed in the March 10, 1994 and May 3, 1994 meetings with EPA and CPDHE. An in-depth analysis and discussion of the physical aspects of the where the OU 3 and background

samples were taken is needed. If the data sets are not comparable from a physical sense (i.e., environmental conditions and flow regimes), a statistically significant difference between site and background will be inconclusive because the test is evaluating the effect of more than one variable. The variable to be tested is the influence of Rocky Flats Plant operations. One will not be able to determine if a difference is due to anthropogenic influences, due to Rocky Flats Plant operations, or due to incomparable physical conditions.

The use of a point-by-point comparison of the OU 3 groundwater data to the upper tolerance limit (UTL) was approved by EPA and CDPHE in the February 14, 1994 meeting. If the comparison is made, no arsenic and beryllium samples exceed the UTL and would, therefore, not qualify as COCs. Also, the groundwater data were not collected to represent the aquifers within OU 3.

Comment 3, Section 3, Page 13. The COC selection process is to be applied by operable unit. DOE's application of the detection frequency criteria is by IHSS. This is incorrect. The entire OU data set should have been considered as a whole.

Response to Comment 3: According to the RFETS COC selection process guidance (CDPHE/EPA/DOE, 1994), COCs are selected using all data, for a particular medium, from an operable-unit. However, for OU 3, the selection of COCs on an OU-wide basis is not appropriate based on spatial, exposure, and different hydrologic and physical processes. Therefore, COCs were selected on an IHSS-basis.

The following points support selection of COCs on an IHSS-by-IHSS basis:

- Performing the concentration-toxicity screen on an IHSS-by-IHSS basis is the *most conservative approach* because it provides opportunity for more compounds to be retained as COCs in the detection-frequency screen. Non-detected data from one IHSS may contribute to lower the overall detection frequency below five percent and eliminate the chemical as a COC. Because of this artifact, a chemical detected greater than five percent of the time in one IHSS, may be eliminated as a COC.
- Performing the concentration-toxicity screen on an IHSS-by-IHSS basis is the *most conservative approach* because it provides opportunity for more compounds to be retained as COCs in the concentration-toxicity screen. For the entire OU data set,

the maximum concentrations for arsenic and beryllium would represent the top 99 percent of the concentration-toxicity score excluding many other chemicals. However, in IHSSs that have smaller maximum concentrations, other chemicals may contribute to the top 99 percent of the concentration-toxicity score and pass the screen (i.e., be retained). As an example, the maximum arsenic sediment concentrations are as follows: 9.4 mg/kg in IHSS 200, 17.7 mg/kg in IHSS 201, and 10.4 mg/kg in IHSS 202. Using the arsenic concentration in IHSS 201 as the maximum OU 3 arsenic concentration, it would contribute almost 50 percent more to the concentration-toxicity score than if the 9.4 mg/kg and 10.4 mg/kg values were used on an individual IHSS basis.

- Each IHSS represents a distinct geographic location that is 1 to 2 miles from the other IHSSs. From a demographic and exposure perspective, different populations would likely frequent different reservoirs and it is not reasonable to aggregate the data in a manner that is inconsistent with exposure patterns. Consideration of population dynamics in the HHRA is discussed in EPA Guidance (EPA, 1989).
- Potential background, RFETS-related, and other anthropogenic sources of metals concentrations are different (e.g., Clear Creek Superfund site, mineral deposits, other commercial, industrial, or agricultural sources). These factors introduce much uncertainty when making comparisons to background by the Gilbert process or the WOE evaluation.
- The source of water for each IHSS are from different watersheds. Mower Reservoir receives approximately 100 percent of its water from the RFETS drainage basin, while Great Western Reservoir receives only 25 to 35 percent from the RFETS drainage basin and Standley Lake receives only 5 to 10 percent from the RFETS drainage basin.
- Exposure scenarios will differ between the IHSSs. For example, Great Western Reservoir may be drained, exposing the surface and potentially the subsurface sediments, while Standley Lake and Mower Reservoir will continue to be used for drinking water and irrigation purposes, respectively. Further, Standley Lake is widely used for recreation while Great Western Reservoir is not. Mower Reservoir is privately owned and used mainly for irrigation.

- The ability to effectively communicate risk to the public will be complicated by OU 3-wide COCs. For example, additional effort would be required to explain to the public that plutonium in Standley Lake is not a problem (i.e., no $^{239/240}\text{Pu}$ activities exceed the 1×10^{-6} PRG), when it has been identified as a COC. COCs identified in each IHSS accomplish the purpose of the COC selection process by focusing the assessment on those chemicals that will contribute significantly to potential risks. Communicating OU 3 risk to the public has been a central theme shared by EPA, CDPHE, and DOE from the outset of the project in 1990. EPA, CDPHE, and DOE could be criticized for taking too broad a view of OU 3.

Comment 4, Section 3, Page 14. Similar to the above comment c, the concentration toxicity screen was applied by IHSS whereas it should have been applied using the entire data set.

Response to Comment 4: This comment has been previously addressed in the Response to Comment 3.

Comment: The above deviations were considered serious enough to warrant an independent analysis of the OU 3 data and selection of COCs by the conventional methodology. The results of this can be summarized as follows:

SURFACE SOIL	Pu-239/240, Am-241
SEDIMENT	As, Be, Pu-239/240, Am-241
GROUNDWATER	As, Be, U-233/234

Response to Comment: The list of COCs developed by EPA is not based on EPA national guidance (EPA, 1989), EPA regional guidance (EPA, 1994), or the approved RFETS COC selection process (CDPHE/EPA/DOE, 1994; Gilbert, 1993; EG&G, 1994a). The chemicals presented by EPA consider only detection frequency, essential nutrients, and toxicity and concentration. EPA's approach ignores the possibility that the reported compounds and their concentrations represent concentrations above background levels (i.e., contamination). This approach does not consider the statement in Section 5.7 of the Risk Assessment Guidance for Superfund (EPA, 1989), "In some cases, a comparison of sample concentrations with background concentrations (e.g., using geometric mean concentrations of the two data sets) is useful for identifying the non-site-related chemicals that are found at or near the site." and "If inorganic chemicals are present at the site at naturally occurring levels, they may be eliminated from the quantitative risk assessment."

In addition, EPA has not provided sufficient technical information on the methods used to select these COCs. The only information provided to DOE was attached to the informal EPA comments received from Bonnie Lavelle on December 6, 1994. The informal comments contain information that contradicts EPA's formal comments on TM 4. As an example, it appears that the handwritten corrections on Table 18 of the informal comments indicate additional steps beyond those described in the text. The table indicates the following steps were performed as part of the EPA COC selection process:

- essential nutrient screen
- frequency of detection screen
- concentration-toxicity screen

However, the handwritten corrections on Table 18 indicate a comparison to PRGs was performed to eliminate additional chemicals as COCs (e.g., barium in groundwater, surface water, and sediments). This action is similar to what was done in TM 4 and contradicts the deficiency noted in Comment 2b on the use of the PRG screen after the concentration-toxicity screen. Although the results of the background comparisons presented in TM 4 for surface soil were used to eliminate uranium-235 as a COC and a spatial analysis argument was used to eliminate uranium-233/234 and uranium-238 as COCs (indicated with handwritten corrections), neither of the screening steps were applied to the other media. Again, this is inconsistent with several of the comments provided by EPA and TM 4. In order to provide a technically verifiable basis for the COC list provided by EPA, DOE requests documentation of the methods and results of EPA's COC selection process.

The concentration-toxicity screen is a zero-sum-like approach, whereby, there will always be compounds retained, no matter how toxic, or at what concentration or whether they represent contamination from RFETS. Significant likely ramifications of performing a HHRA using the above information include:

- Consideration of groundwater east of Standley Lake and Great Western Reservoir as impacted by Rocky Flats has no scientific basis in the regional hydrogeologic regime. Suggesting that contamination from Rocky Flats has migrated to these environs is not reasonable and could mislead decision makers and the public.
- Consideration of arsenic and beryllium as the contaminants of concerns, and, therefore, public health threats, also misleads decision makers and the public. Any discrepancy in

their concentrations suggesting other than natural occurrence can be attributed to subtle variation in the physical and chemical environment and not to a release of contaminants in the environment.

Comment: Also, those chemicals without toxicity factors that are present above background and at greater than 5% frequency of detection should be identified for each medium. The potential impact on the human health risk must be addressed qualitatively in the human health risk assessment. The following chemicals are in this category:

SEDIMENT	aluminum, cesium, cobalt, lead, lithium, silicon, thallium
-----------------	--

GROUNDWATER	aluminum, cobalt, iron, lead, lithium, silicon
--------------------	--

SURFACE WATER	aluminum, cobalt, lead, lithium, silicon
----------------------	--

Response to Comment: All of these compounds are naturally occurring and are ubiquitous. There is no rationale for their inclusion as COCs other than the fact that EPA has not published toxicity information for them. A qualitative discussion of this information regarding the above chemicals will be included in the HHRA.

SPECIFIC COMMENTS

Specific Comment 1, Section 2.2.2, Page 4. Differences in quality assurance (QA) procedures between the 1983-1984 data and more recently collected data are discussed in the second paragraph of this section. DOE conducted a statistical comparison to determine if the two data sets could be combined. DOE concluded that they were similar and could be combined. However, it is not clear whether the more recent samples were collected from the same sample locations as the 1983-1984 samples. If sampling locations were not the same, then the statistical tests are actually evaluating differences between locations or sampling methodology, as well as other potential differences. Additionally, because different QA procedures were used for 1983-1984 data, sample quantitation limits may not be comparable between the two data sets. For example, if the two data sets have different sample quantitation limits; they cannot be directly compared. These complications should be addressed in this section, and other sections which describe combining data.

Response to Specific Comment 1: Many of the RFI/RI sediment samples were collected to correspond to known 1983 and 1984 sample locations (See Figures 2-2 and 2-3 in TM 4). Two pair-wise statistical tests were performed--only those locations with both sets of samples (RFI/RI samples and 1983 and 1984 samples) were included in the analysis. The paired location sample numbers are identified in memorandum by S. Blake/CH2M HILL, dated November 10, 1993 (included in Appendix A of TM 4 (DOE, 1994)). The paired analyses performed (a paired t-test and a Sign test) tests the hypothesis of whether the mean of the differences at each sample location are significantly different from zero. This type of analysis takes into account differences between sample locations and, therefore, differences between sampling locations is not an issue. Both statistical tests show no significant difference in the 1983/84 data and the Standley Lake data at a 95 percent confidence. However, the mean and median ^{239/240}Pu activity level of the 1983/84 Great Western Reservoir data is higher than the 1992 RFI/RI data. Based on these results, the 1983/84 data was combined with the RFI/RI data.

An assessment on the reported sample quantitation limits (SQLs) was not performed. Differences between the SQLs of the data sets may impact the results of the statistical comparison tests. Furthermore, as stated in the RFI/RI Work Plan for OU 3 (DOE, 1992), extensive QA/QC information is not readily available for the 1983/1984 sediment samples and, after extensive investigation, the locations of several sample locations are not known. While these data do have QA/QC issues, one of the purposes of the OU 3 field program was to confirm historical data and use it to the maximum extent possible. Also, it was determined that the inclusion of the 1983/1984 sediment data for Great Western Reservoir would be conservative since these data have higher values than the RFI/RI data.

Specific Comment 2, Figure 3-4, Page 8 or 9 of Section 3. This figure presents the background comparison methodology. According to the Gilbert methodology, an additional step should be included in the flow chart before the slippage test. The slippage test should be used if the highest datum is a detect. If not, then the next step should be to determine if there are less than 20 percent nondetected samples in the site and background, and whether the site and background data are normally distributed. The figure should be corrected to include this step.

Response to Specific Comment 2: We agree with the comment. An additional step should have been identified in the flow chart, occurring before the slippage test, to check whether the

largest background value is a detect. The slippage test is used only if the largest background data value is a detect. Although this step was inadvertently left off the flowchart, the background comparison methodology was employed correctly for the OU 3 surface soil data set. The flow chart should correspond to the approved flowchart in the Strawman guidance (REF).

Specific Comment 3, Section 3.5, Page 14. This section describes the CTS screen used to select COCs and Appendix D presents the CTS tables.

Although the description in Section 3.5 accurately explains how to conduct a CTS, the CTS tables do not present the information necessary to easily verify the results of the assessment. The tables in Appendix D should be revised to include the maximum detected concentration and toxicity value used for each chemical, as well as the chemical risk factor, total risk factor, and the ratio of each individual chemical risk factor to the total risk factor.

Response to Specific Comment 3: New concentration-toxicity information tables for the RFI/RI Report containing the additional information described in the comment will be provided on request to EPA.

Specific Comment 4, Section 3, Page 16. The weight of evidence evaluations fall short of EPA's expectations because no criteria were established or apparently applied to discriminate appropriate literature values from inappropriate ones. At a minimum, we expected some consideration of the geologic materials comprising the sediment background locations compared to OU 3 conditions, an evaluation of flow conditions, an evaluation of the uncertainty in each estimate of "background" from the literature (i.e., sample size, sampling methods, QA/QC considerations) and an evaluation of location of the "background" samples relative to anthropogenic sources of contamination. Instead of providing useful information, it introduces much uncertainty to the COC selection process. A comparison to other contaminated Superfund sites was also done with the OU 3 data. This has no relevance to the question of whether sediments, surface water, and groundwater in OU 3 contain chemicals above background concentrations.

Response to Specific Comment 4: All available information was gathered, beginning with information from the Background Geochemical Characterization Report (BGCR) (DOE, 1993) and from areas along the Front Range. These data were supplemented by other Colorado and

national data sets. A systematic process was used to evaluate benchmark data sets. The only data sets not included were some arsenic and beryllium background concentrations from the U.S. These concentrations were at similar levels as the arsenic and beryllium concentrations already presented. No other data sets found during research were eliminated.

The following sources, in order of preference, were accessed:

- Data from the BGCR
- Metro Denver data
- Front Range data
- Colorado data
- US data
- World data

Pertinent observations of these data sets include:

- Surface water data is from freshwater sources
- Majority of data presented in TM 4 are from US sources
- Some of the data is from the front range of the Rocky Mountains within 20 miles of OU 3
- Data published in scientific journals, books, or other scientific literature
- Data collected by agencies responsible for maintenance, operation, organization, etc. of a land use (contamination)
- The supporting information to perform a rigorous QA/QC evaluation was not available; however, most of the reviewed data were obtained from published scientific sources or organizations (e.g., USGS) and would not be expected to have been published without proper QA/QC.

Uncertainty does exist in the quality and usability of the benchmark data, but realizing this uncertainty when using these data in the comparisons, and combining the conclusion reached from these comparisons with the other WOE evaluation steps bolsters the conclusion that these data represent general background conditions as do the OU 3 data. These data sets have been published in scientific journals, books, or other scientific literature, and media which require a high level of QA/QC.

The WOE evaluation considers concentration levels for each chemical from each IHSS. Since each IHSS would likely receive contamination at different times from different release events, IHSS-specific contamination should be apparent in the evaluation. Based on the concentrations seen in these IHSSs, arsenic and beryllium concentrations are within the background ranges. Additionally, considering that 90 percent of the water going into Standley is from Clear Creek and only 25 to 35 percent of the water flowing into Great Western Reservoir is from the North and South Walnut Creeks, and approximately 100 percent of the water flowing into Mower Reservoir is from RFETS, the concentrations are remarkably similar; further supporting the determination that these metals are within the background ranges.

As an example, the arithmetic mean for arsenic concentrations in the stream sediments (creeks and drainages) of IHSSs 200, 201, and 202 are 5.3, 4.8, and 4.9 mg/kg, respectively, and in the reservoir surface sediments (lakes) are 4.9, 6.9, and 5.1 mg/kg, respectively. The arsenic concentration ranges are from 3.7 to 9.4, 2.2 to 7.8, and 3.0 to 6.8 mg/kg for stream sediments in IHSSs 200, 201, and 202, respectively; and 2.6 to 9.4, 1.2 to 17.7, and 2.2 to 10.4 mg/kg for reservoir surface sediments in IHSSs 200, 201, and 202, respectively. These comparisons suggest comparability, not divergence, in the low part per million range. Comparing to the stream sediment data reveals similar concentration levels. The BGCR arsenic range in stream sediments is 0.2 to 17.3 mg/kg with a mean of 2.4 mg/kg and the BGCR beryllium range is 0.15 to 1.3 mg/kg with a mean of 0.7 mg/kg. An arsenic concentration of 2.4 mg/kg translates to a 6×10^{-6} risk and a beryllium concentration of 0.7 mg/kg translates to a 5×10^{-6} risk (the risks are based on residential exposure parameters from the Programmatic Preliminary Remediation Goals document (DOE, 1994a)). If arsenic and beryllium are considered COCs, then based on the analytical data all IHSSs--Great Western Reservoir, Standley Lake, and Mower Reservoir--and the entire area of each IHSS (i.e., every part of each stream and reservoir) are contaminated.

Following EPA's reasoning that the background data are not comparable to OU 3 conditions, based on a non-statistical comparison, the areas where the BGCR data were collected would then be considered contaminated, as would the Cherry Creek reservoir, the Rocky Mountain National Park lakes, the background stream sediments for the Lowry Landfill Superfund site,

and the Great Lakes, Lake Adirondack, and Lake Michigan. Clearly, not all these areas have been contaminated to levels exceeding background. Rather, this illustration exemplifies the commonly observed natural variation in the physical environment. Subtle differences or variations do not normally indicate pollution as a source of variation. The concentration data used to represent the benchmark (background) levels is very consistent (see the bar graphs in TM 4, for example, Figure 5-1). There are no apparent large fluctuations and only the data identified as site contamination from Superfund sites are much greater than all other concentrations. Comparison to concentration levels from Superfund sites illustrates the typically encountered chemical concentration levels found at hazardous waste sites. Based on experience, the levels of contamination from the release of hazardous substances is not subtle and the identity of released contaminants is normally indicated by appreciable increases above natural and wide-spread anthropogenic levels.

Additionally, the Gilbert process (Gilbert, 1993; DOE, 1994c; EG&G, 1994a; EPA, 1993) includes three professional judgment guidelines that are used in conjunction with the statistical tests and do not require the use of benchmark data:

- Spatial distribution arguments
- Temporal distribution arguments
- Pattern recognition concepts

Spatial distribution argument and temporal distribution arguments are an integral part of the WOE evaluation and agree with the results of a comparison to benchmark data.

Specific Comment 5, Section 3.10, Page 29. This section describes how Phase 1 Historical Public Exposure Studies were used to confirm chemicals as COCs. The purpose of the historical studies was not to support risk assessment or COC selection for OU 3. As stated in this section, more than 8,000 chemicals were identified as having been used at the Rocky Flats site, but “the list was reduced to those chemicals that were most likely to have posed an offsite human health hazard under routine historical plant operations.” The focus of historical investigation was not to select COCs. For example, Table 3-3 lists the materials of concern by the RFETS health studies. Most of these chemicals were eliminated as COCs through the selection process used in TM 4.

Response to Specific Comment 5: The materials of concern presented in the Phase I Historical Public Exposure Studies support the results of the WOE evaluation and were presented in TM 4 for information purposes only. Although, the purpose of the Phase I studies was not to support OU 3 activities, the methods used to identify the materials of concern are generally more rigorous for identifying potential sources of contamination to the offsite area than source definition methods related to the OU 3 RFI/RI activities. According to ChemRisk, speaking to the Phase I Health Studies:

"The initial tasks (including the inventory and selection of COCs) deal with the review and compilation of historical information for the purposes of selecting specific radionuclides and chemicals that warrant detailed study as well as accidents or incidents that may have affected the offsite public." -- ES&T Vol. 26, No. 7

In the study ChemRisk employed a WOE method which does consider the toxicological properties of the materials used at Rocky Flats and also considers accidents or other incidents beyond routine plant operations. Only the chemicals classified as materials of concern that were used in sufficient quantities or were released during any routine or non-routine event to be considered a contaminant source. For example, according to the study, arsenic was not used during plant operations and there were no known releases of arsenic. Additionally, beryllium, a material of concern, is a potential source. But upon further analysis of the concentration data within OU 3, something not performed as part of the Phase I Health Studies, beryllium is not a COC.

Based on the above information, it is reasonable and scientifically appropriate to use the results of these methods to support risk assessments and COC selection tasks related to Rocky Flats.

Specific Comment 6, Section 4, Page 4. Table 4-2 is inconsistent with the information in Appendix B. Appendix B indicates that plutonium activity in soils is not normally distributed. Therefore, the t-test is not a valid statistical test.

Response to Specific Comment 6: We disagree with the EPA comment that Table 4-2 is inconsistent with Table B-1 in Appendix B regarding the t-test. The criteria for performing the t-test is defined in the RFP Guidance Document, Revision 0, Draft B (DOE, 1994a) and the approved Gilbert guidance document (i.e., the "Strawman") (DOE, 1993a; EPA, 1993) as follows:

“IF A) EITHER both background and OU data contain at least 20 data points, OR both distributions are normally distributed
AND B) less than 20% of the background and OU data are classified as non-detects,
THEN use the t-test.”

None of the data analyzed follow a normal (or log normal) distribution in both the OU 3 and background. However, Table B-1 shows $^{239/240}\text{Pu}$ contains at least 20 data points in both OU 3 and background (OU 3 -109, background-20). Therefore, according to the criteria above, a t-test should be performed and the resulting p-value is shown in Table 4-2.

Specific Comment 7, Appendix G. This appendix provides probability plots used in the weight-of-evidence evaluation. Radium-226 is not discussed in the text although a probability plot has been provided for it. Radium-226 should be discussed in this appendix.

Response to Specific Comment 7: The probability plot was inadvertently included in Appendix G. Radium-226 is eliminated in the PRG-screen and, therefore, a WOE evaluation presentation is not needed.